

“Big Signal Processing for Multi-Aspect Data Mining”



Thursday, January 14, 2016
3:00pm-4:00pm, IS Building, 3rd floor
135 North Bellefield Ave Pittsburgh, PA

What does a person’s brain activity look like when they read the word apple? How does it differ from the activity of the same (or even a different person) when reading about an airplane? How can we identify parts of the human brain that are active for different semantic concepts? On a seemingly unrelated setting, how can we model and mine the knowledge on web (e.g., subject-verb-object triplets), to find hidden patterns and missing links? My proposed answer to both problems (and many more) is through bridging signal processing and large-scale multi-aspect data mining.

In addition to the above Neurosemantics application, multi-aspect mining appears in numerous applications such as mining knowledge on the web (where different aspects of the data include entities in a knowledgebase and the links between them or search engine results for those entities) and multi-aspect graph mining (with the example of multi-view social networks) where we observe social interactions of people under different means of communication, and we use all views/aspects of the communication to extract more accurate communities.

**Evangelos
Papalexakis**

The main thesis of my work is that many real-world problems, such as the aforementioned, benefit from jointly modeling and analyzing the multi-aspect data associated with the underlying phenomenon we seek to uncover. In my research, I develop scalable and interpretable algorithms for mining big multi-aspect data, with emphasis on tensor factorization. In this talk, I will discuss multi-aspect data applications, focusing on Neurosemantics, and present my algorithmic work on scaling up tensor factorization by two orders of magnitude and assessing the quality of the results. I conclude with my future vision on bridging Signal Processing and Data Science for real-world applications.

Bio: Evangelos (Vagelis) Papalexakis is a PhD candidate at the School of Computer Science at Carnegie Mellon University (CMU). Prior to joining CMU, he obtained his diploma and MSc in Electronic & Computer Engineering at the Technical University of Crete, in Greece.

His research interests span the fields of data science, data mining, signal processing, and machine learning. His research involves designing scalable algorithms for mining large multi-aspect datasets, with specific emphasis on tensor factorization models, and applying those algorithms to a variety of real world, multi-aspect data problems.

His work has appeared in KDD, ICDM, SDM, ECML-PKDD, WWW, PAKDD, ICDE, ICASSP, IEEE Transactions of Signal Processing, and ACM TKDD. He has a best student paper award at PAKDD'14, finalist best papers for WWW Web Science Track '15, SDM'14 and ASONAM'13, and he was a finalist for the Microsoft PhD Fellowship and the Facebook PhD Fellowship. Besides his academic experience at CMU, he has industry research experience working at Microsoft Research Silicon Valley during the summers of 2013 and 2014 and Google Research during the summer of 2015.